Bird Biodiversity Information System in Indonesia Based on Web Semantic

[1]Futty Pratiwi
[2]Adang Suhendra
[3]Arrummaisha Adrifina

[1]Gunadarma University(futtykyut@staff.gunadarma.ac.id)
[2]Gunadarma University(adang@staff.gunadarma.ac.id)
[3]Gunadarma University(arroem@staff.gunadarma.ac.id)

Abstract

Web technology has developed rapidly, began with Web 1.0, until the latest Web 3.0. The concept of Web 3.0 is more focused on semantic web which refers to the ability of computers application to understand human language better, not only the standard language of the users but also to more complex language.

Biodiversity is being intensively discussed by researchers in the world due to climate change in earth. These researchers need data from various parties regarding the biodiversity, in this case is about birds, to be able to anticipate many bird conservation problems due to changes in the earth conditions. But often, a variety information on bird biodiversity web generally only serve as information and does not provide data that can be exploited by others.

This Bird Biodiversity Information System based on Semantic Web not only has function to entry new bird data, but also able to convert the data into RDF (Resource Description Framework) form and shows the results of queries based on the keyword given by user. RDF is a language for processing metadata, where metadata can be encoded, exchanged and used in the web. This system is made by using combination of software that supports the development of a dynamic and interactive website which is PHP, MySQL, RDF API for PHP and Apache Server. Keywords: Semantic, Bird, RDF, Metadata

1 Introduction

Now, Website is a necessary part in modern human life. Not only for online transaction, but also for spreading the information. Web technology also has developed, starting from Web 1.0 to the lastest version of Web technology Web 3.0. The concept of web 3.0 is more focused to semantic web which refers to computer capability in understand human language.

Biodiversity is being intensively discussed by biologic researchers in the world due to changes in the earth condition. the biologic researchers need data about biodiversity from various parties. Often, a variety of information on bird biodiversity site only serves as information and does not provide data that can be utilized other parties, so there are many sites of bird biodiversity but has different information structure.

To overcome the problem, we needed an information semantic web-based system bird biodiversity that has functions not only entry new bird data, but also able to convert data into RDF format. RDF is an XML-based formal language. RDF is a foundation for processing metadata, where metadata can be encoded in the web, exchanged and used. Defining the semantics of RDF can do more understandable explanation for the engine to facilitate data exchange.

Semantic Web as a new generation of web technology is very promising in building a biodiversity web semantic service. Researchers who are in separate locations can reported the discovery of new biological diversity through the Internet, then the data will be converted into metadata format. We hope the existence of metadata can facilitate the dissemination, discovery and using of bird biodiversity information. For the first step, the authors chose Taman Mini Indonesia Indah Bird Park and Perhimpunan Pelestarian Burung Liar di Indonesia as a place to get birds data collection.

2 Research Method

2.1 Analyze

This research begins with the analysis process. Analysis was done so that the problems previously discussed in more detail in mind so it can find an appropriate solution. Analysis based on literature study to gain basic theories as a source of reference in doing research. Information and literature related to this problem is obtained from:

1. Electronic books and journals written in the form of semantic web related.
2. Internet, such as examples of semantic web application.
3. Sources of other information, the explanation given in the form lecturer and fellow students.

2.2 Design

Furthermore, the system model is designed using UML(Unified Modelling Language) and the interface of the system according to the required functions.

2.3 Implementation

Finally, the design has been created and then implemented into the programming language PHP, RDF, and RAP (RDF API for PHP).

3 Biodiversity Classification

Classification of living things based on similarities and differences of living things possessed traits such as body shape, or function of the appliance body [4]. Living beings that possess the same characteristics are grouped into one class. Examples of classification of living beings is based on body size, neighborhood life, based on merit, based on the type of food and based on the binomial naming procedure [2].

An international organization named GBIF (Global Biodiversity Information Facility) makes a biodiversity classification based on taxonomy structure and datasets[1]. The structure is:

Common Name, or known as English Name
Kingdom, based on biology taxonomy
Phylum, based on biology taxonomy
Class, based on biology taxonomy
Order, based on biology taxonomy
Family, based on biology taxonomy
Genus, based on biology taxonomy
Species, based on biology taxonomy

4 Data Standardization

Of the various classifications, this work try to create a new data structure standardization that is based on the availability of preliminary data from the Bird Park in TMII, Perhimpunan Pelestarian Burung Liar Indonesia[3] and Global Biodiversity Information Facility [1] as following:

Indonesian Name
Common Name, or known as English Name
Kingdom, based on biology taxonomy
Phylum, based on biology taxonomy
Class, based on biology taxonomy
Order, based on biology taxonomy
Family, based on biology taxonomy
Genus, based on biology taxonomy
Species, based on biology taxonomy
Color
Size, in centimeters
Region
Elevation, in meters
Observer

5 System Modeling

In designing this system, UML modeling tools is used. This system will be described in 3 kinds of UML, thry are: use case diagrams, activity diagrams and sequence diagrams.

5.1 Use Case Diagram

This diagram shows the relationship between the systems with users outside the system. The users are named user and admin.

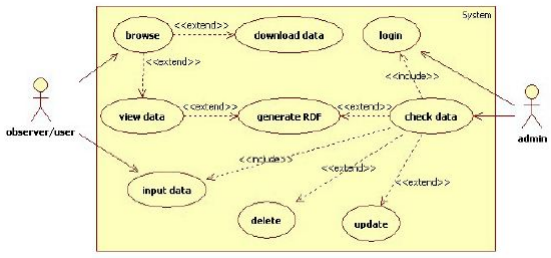

Figure 1: Use Case Diagram

5.2 Activity Diagram

Activity diagram is a diagram for illustrating the workflow in the system. This diagram showed activity that causes an object is in a particular state.

Table 1: Narrative Specification for Input Data Activity

Use case	**Activity**
User choose menu input	navigate to input data
System display input form	built input form
User input data	enter and submit data
System validate data	form validation
System respond for wrong input	show info failed
System respond for valid input	show info success

Red-circled symbol indicates the initial state while the circled-symbol wrapped by the striped circle indicates the end of the state. Activity diagram that authors create include input data, browse, and check data use case.

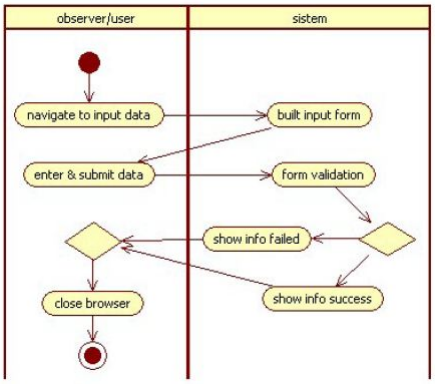

Figure 2: Input Data Activity

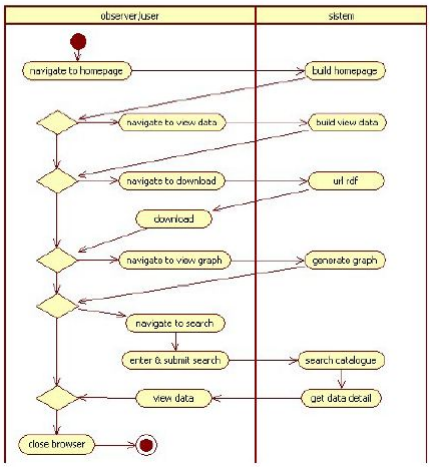

Figure 3: Browse Data Activity

Table 2: Narrative Specification for Browse Data Activity

Use Case	Activity
User choose menu browse	navigate to browse data
System display home	Build Homepage
User choose menu view	navigate to view data
System display home	built view data
User choose menu download	navigate to download
System display download link	show RDF url
User choose sub menu download RDF	download
User choose menu view graph	navigate to view graph
System display graph	generate graph
User choose menu search	navigate to search
User enter keyword	Enter & submit search
System do searching	search catalogue
System display query	get data detail
User view data	view data

Table 3: Narrative Specification for Check data Activity

Use case	Activity
Admin enter username and password to login	navigate to login form
System validate username and password	login validation
System respond for invalid entry	info failed
System respond for valid entry	view list data
Admin change data	update data
System display success message	data updated
Admin delete data	delete data
System display success message	data deleted
Admin generate new data entry	generate to RDF

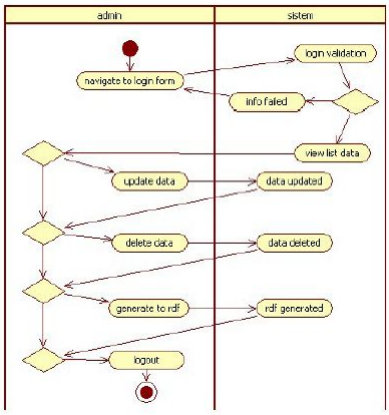

Figure 4: Check Data Activity

5.3 Sequence Diagram

Sequence diagrams describe interactions between objects within and around the systems as a message that described based on time. Sequence diagrams consist of vertical dimension (Time) and horizontal dimension (related objects). Sequence diagram is used to describe a scenario or series of steps undertaken as a response from an event to produce a given output.

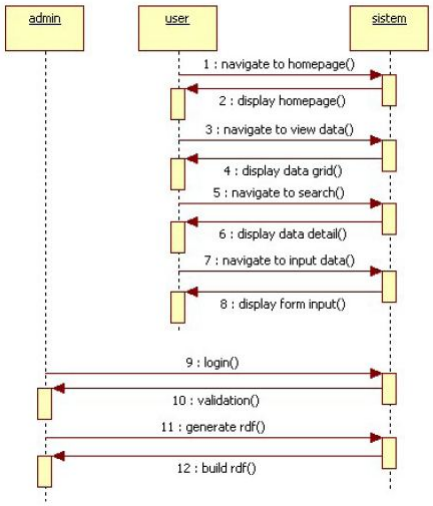

Figure 5: Sequence Diagram

5.4 Service Oriented Architecture

Figure 6 shows the architecture of the information system which author develops. The information system includes entering data and automatically converted to RDF format and query based on keywords using SPARQL. SPARQL(**S**PARQL **P**rotocol and **R**DF **Q**uery **L**anguage) is a query language for RDF. It was standardized by World Web Consortium, and is aconsidered a key semantic web technology.

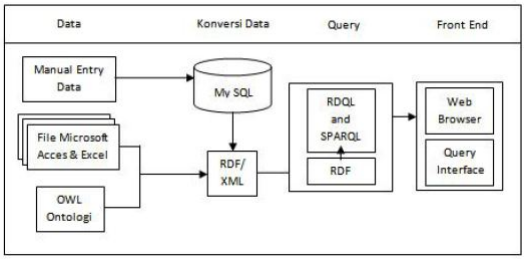

Figure 6: Service Oriented Architecture

6 RDF

In this system, the author generates two type of RDF namely RDF Structure and RDF Graph.

6.1 Bird RDF Structure

Data representation format for the semantic web is RDF (Resource Description Framework). RDF is a standard model for data exchange on the web. Here is the structure of RDF Bird to be displayed on web-based information system the authors make.

<? xml version="1.0" encoding="UTF-8" ?> <rdf:RDF xml:base="http://bird.gunadarma.ac.id/burung#" xmlns:Description="http://loc alhost/futty/Animal.owl#" <rdf:Description rdf:about="http://loc alhost/futty/Birds.owl#NorthernCassowary"> <Description:type>Bird</Description: type> <Description:hasCommonName>Northern Cassowary</Description:hasCommonName> <Description:hasIndonesianName> Kasuari Gelambir Tunggal </Description:hasIndonesianName> <Description:hasKingdom>Animalia </Description:hasKingdom> <Description:hasPhylum>Chordata </Description:hasPhylum> <Description:hasClass>Aves </Description:hasClass> <Description:hasOrder>Casuariiformes </Description:hasOrder> <Description:hasFamily>Casuariidae </Description:hasFamily> <Description:hasGenus>Casuarius </Description:hasGenus> <Description:hasSpecies>Casuarius unappendiculatus </Description:hasSpecies> <Description:hasRegion>Papua </Description:hasRegion> <Description:hasColor>Black </Description:hasColor> <Description:hasObserver>anonim </Description:hasObserver> <Description:hasSize>100 </Description:hasSize> <Description:hasElevation>700 </Description:hasElevation> </rdf:Description> </rdf:RDF>

Due to data flexibility, then the possibility of change RDF structure above is very large later on.

6.2 Bird RDF Graph

The visual appearance of RDF structures is in the form of an RDF graph consisting of triple subject-predicate-object. Subjects were shown on the graph is a URI from Bird. While objects were shown on the graph is the URI of the range of object properties. Link between them is a Predicate that displays

the URI of the object properties. Graph will be generated in accordance with the existing RDF file contents. So there will be an extensive data network when the database has a large numbers. There are 8 major components for the Bird, they are: CommonName, IndonesianName, Kingdom, Phylum, Class, Order, Genus, Color, Size, Observer, Region, Elevation and Family. Each component serves as Object. While acting as a subject is the namespace of the Birds. Between subject and object are connected by the predicate. Predicate includes object properties as follows: hasCommonName, hasIndonesianName, hasKingdom, hasPhylum, hasOrder, hasGenur, hasSpecies, hasColor, hasSize, hasObserver, hasRegion, hasElevation and hasFamily. Here is a RDF graph that generated by Bird RDF file.

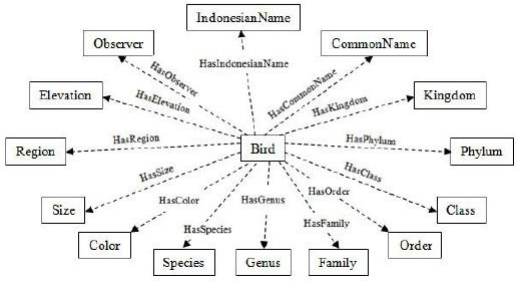

Figure 7: RDF Graph

7 Data Sample

After standards are determined and the structure of RDF data is made, the next step is to determine sample data to be used in this information system. Preliminary data obtained from Taman Burung Taman Mini Indonesia Indah and Perhimpunan Pelestarian Burung Liar Indonesia. Here is a sample of data used in the semantics web-based information system.

8 Interface Design

Interface design is a desired view estimation of the site and followed by explanation of each section. On this web-based system, there are two sides of view, namely regular user page views and admin page views. The user page can be used by anyone to enter a new bird data and download the RDF forms. While the admin page can only be used by certain people who have authorization to confirm the new data entered by users.

Table 4: Sample of Bird Data

Kategory	Keterangan
Common Name	Sunda Coucal
Indonesian Name	Bubut Jawa
Kingdom	Animalia
Phylum	Chordata
Class	Aves
Order	Cuculiformes
Family	Cuculidae
Genus	Centropus
Species	Centropus nigrorufus
Color	Hitam
Size	46
Region	Jawa
Elevation	0
Observer	anonim

8.1 User Interface Design

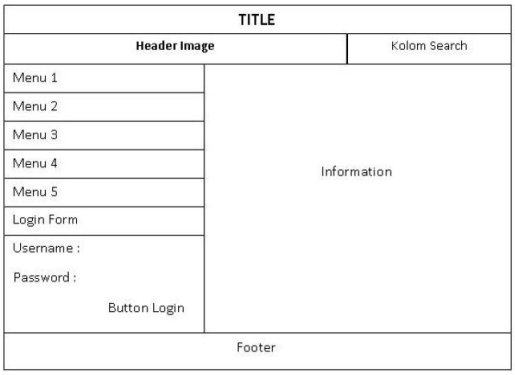

Figure 8: User Interface Design

8.2 Admin Page Interface Design

Here is an interface design for the admin side. There are three menus namely Query, List, and Generate Data File RDF. Admin uses Query menu to perform queries against existing data. The queries are based on the namespace of each subject, predicate object.

The second menu is a Data List. In this menu, the admin can organize existing data. Admin can do the editing and deletion of data. This is necessary if data changes in the future.

The next menu is Generate RDF file, serves to confirm the data entered by the user. If the admin has not done this command, then the new data user entered will not be displayed on the RDF file and data grid.

8.3 Navigation Structure

Navigation structure is used to describe all of the functions in system. The navigation structure in this system consists of navigation structure for the user and admin level, because user and admin has different access rights interacting with the system. Here is a picture of navigation structure used in this system.

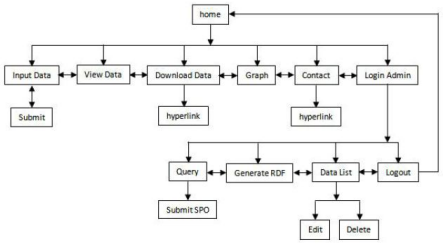

Figure 9: Navigation Structure

9 Conclusion and Suggestion

After conducting several tests, it can be concluded that The Bird Biodiversity Information System in Indonesia Based on Semantic Web that already developed can already be used in public. Because this system has function of entering new birds data which allows researchers to enter new birds data.

The data will automatically be converted to RDF format. Besides, the system also able to displays the RDF structure based on graph and conduct data searching based on key words that is entered by users. The system interface is made as simple as possible so the user can use it easily.

This web-based system is still far from perfect. Therefore the improvement is much needed. Like there is no integration test with the web services.

Suggestion and future work for this Bird Biodiversity Information System in Indonesia Based on Semantic Web, expected that a time services will be added so that can be integrated with the other bird biodiversity information system in international scale.

References

[1] GBIF. Global biodiversity informatics facility. http://www.gbif.org, 2010.

[2] http://id.wikipedia.org/wiki/Taksonomi. Taksonomi, 2010.

[3] Perhimpunan Pelestarian Burung Liar Indonesia. Data burung indonesia, 2009.

[4] C. Leveque & J. Mounolou. *Biodiversity*. John Willey, New York, 2003.